Normalizing War *

Miles D. Williams *Denison University*

This document provides an introduction to R Markdown, argues for its benefits, and presents a sample manuscript template intended for an academic audience. I include basic syntax to R Markdown and a minimal working example of how the analysis itself can be conducted within R with the `knitr` package.

Keywords: pandoc, r markdown, knitr

Introduction

How should we best model the deadliness of international war?[1] The prevailing wisdom holds that war deaths follow the power-law, a distributional form characterizing phenomena with many small events but a few very extreme ones. This view, however, is more often assumed than formally tested. This study conducts a formal test of war's correspondence with the power-law versus a plausible alternative: the log-normal.

Analysis of per capita battle deaths from 95 international wars fought between 1816 and 2007 shows that neither the power-law nor the log-normal distribution can be rejected as a good fit for the data. Further, in a head-to-head test, one is not found to be statistically better than the other. However, while there is no formal basis for favoring one model, the log-normal distribution confers the advantage of not having to trim the data to optimally fit the distribution. While the power-law only is a good fit for the 45 most intense wars, the log-normal is a good fit for all 95 in the data.

Even as the data provide no formal basis for picking one distribution over the other, the choice has the potential to alter statistical inferences. Past research has sought to identify whether the so-called "long peace" in the second half of the 20th century reflects a statistically detectable change in the trend of war's deadliness. Past research has yielded

*Replication files are available on the author's Github account (http://github.com/milesdwilliams15). **Current version**: July 25, 2023; **Corresponding author**: williamsmd@denison.edu.

[1] "War" and "international war" are used interchangeably throughout the paper.

conflicting results. Much of the disagreement may have to do with the way war deaths are differently modeled across studies. Using the 1950 change point recently identified by X, the statistical significance of pre- and post-1950 differences in war deaths is tested using both power-law and log-normal fits for the data.

The choice of model is found to influence conclusions. While a statistically significant difference is detected using a log-normal fit, no such difference is detected using the power-law. This raises questions that need to be addressed in future research. That model choice substantively influences conclusions from statistical inference points to a major limitation in current quantitative study of war fatalities. Unfortunately, 95 wars may not be enough points of observation to resolve this issue. Until more data become available, we are left with little formal justification for favoring one distributional form over the other.

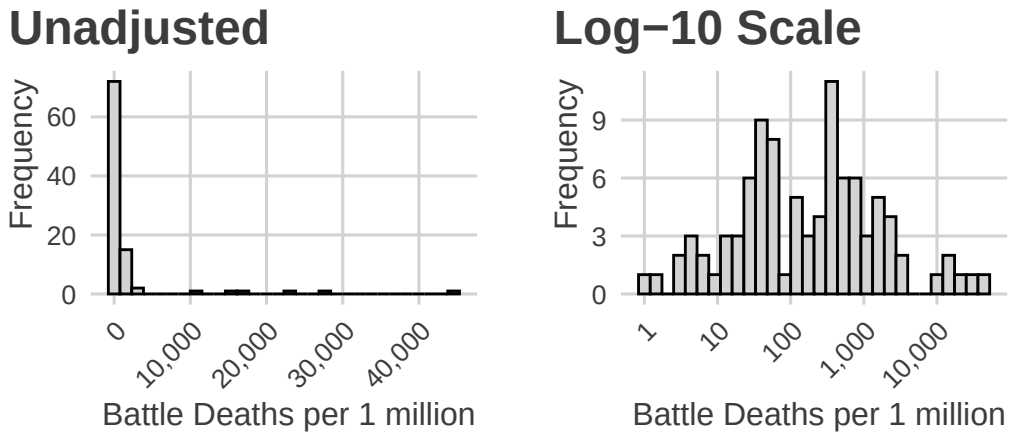

Figure 1: The distribution of battle deaths per belligerent populations in millions for all international wars, 1816-2007.

The software package developed by [x] makes it possible to test and compare four different distributional forms for continuous data.

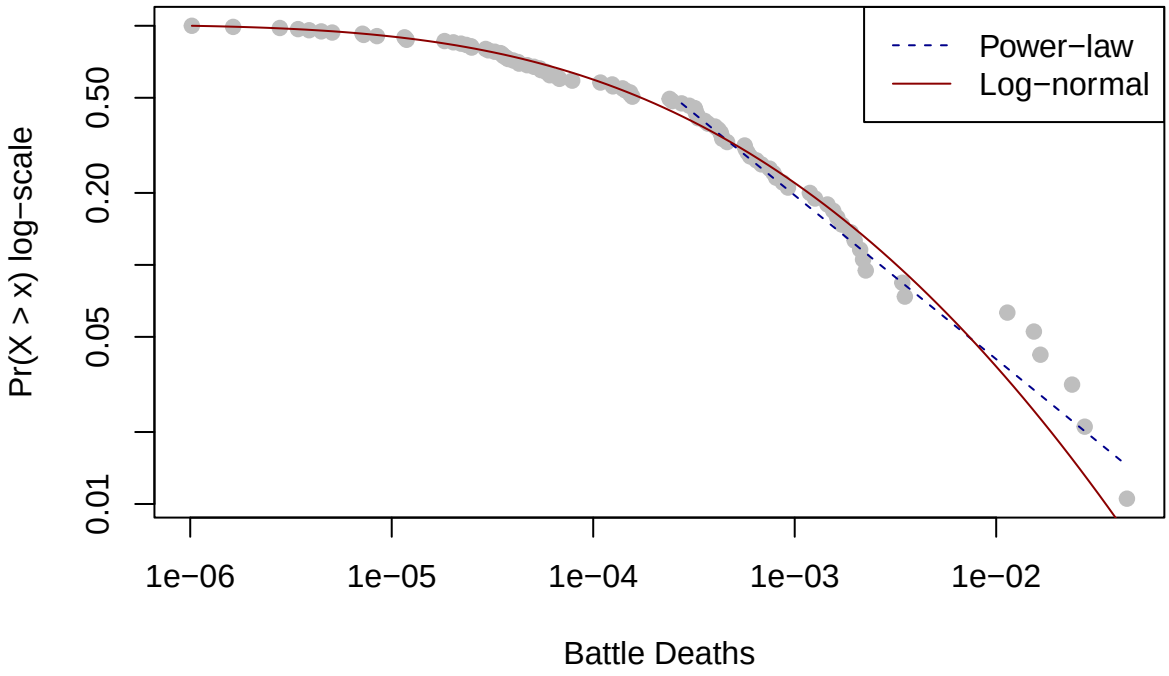

Figure 2: Comparison of the power-law and log-normal fits for the data.

Table 1: KS-test results

Hypotheses	p-value
Can we reject the power-law?	0.850
Can we reject the log-normal?	0.790
Can we favor one over the other?	0.604

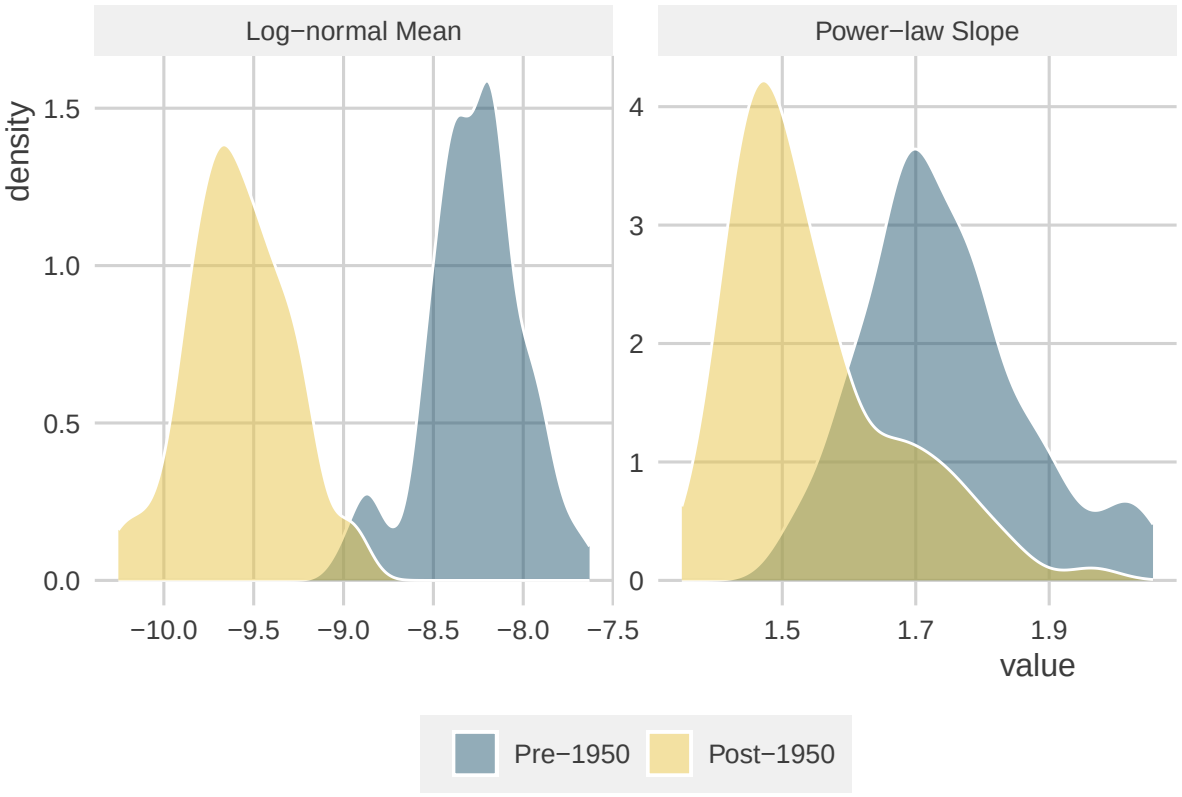

Figure 3: Bootstrap distributions of model parameters pre- and post-1950.